FEBS Letters 404 (1997) 148–152 FEBS 18259

A novel delta-endotoxin gene *cryIM* from *Bacillus thuringiensis* ssp. *wuhanensis*

A.B. Shevelev[a,*], Ya.N. Kogan[b], A.M. Bushueva[b], E.J. Voronina[b], D.V. Rebrikov[b], S.I. Novikova[d], G.G. Chestukhina[a], V. Kuvshinov[c], E. Pehu[c], V.M. Stepanov[a]

[a] *Lab. of Protein Chemistry, Institute of Microbial Genetics (GNIIGenetika), 1st. Dorozhny, 1, Moscow 113545, Russian Federation*
[b] *Molecular Biology Dept., Biological Faculty, Moscow State University, Moscow, Russian Federation*
[c] *Plant Production Dept., Helsinki University, Viikki, Helsinki, Finland*
[d] *Molecular Biophysics Dept., Moscow Inst. of Physics and Technology, Moscow, Russian Federation*

Received 10 January 1997

Abstract A new *cryI*-related sequence designated *cryIM* was cloned using an immunoscreening technique from ssp. *wuhanensis* of *Bacillus thuringiensis* (*BT*), previously reported to produce multiple Cry proteins [Chestukhina et al. (1994) Can. J. Microbiol. 240, 1026–1034]. Analysis of the *cryIM* nucleotide sequence revealed an ORF, BTII-type promoter-like sequence, peculiar for such genes, a translation initiation element and a putative transcription terminator. Nevertheless, its product was not previously found in the crystals of the host strain [Chestukhina et al. (1994) Can. J. Microbiol. 240, 1026–1034] which shows its weak or absent natural expression. The amino acid sequence of 1151 residues encoded by the continuous reading frame of *cryIM* is not identical but is essentially similar to the other δ-endotoxins of the CryI class. An IS231-like sequence was found 400 bp downstream of the *cryIM* stop codon and a fragment of the *cryIAb* gene was located 3 kb upstream of its initiator codon in the same orientation. Artificial expression of the cloned gene in *E. coli* under the control of the *lacZ* promoter allowed us to obtain its hypothetical protein product.

Key words: CryIM; Delta-endotoxin; Primary structure; *Bacillus thuringiensis*; IS231; CryIAb

1. Introduction

Many strains of *Bacillus thuringiensis* (*BT*) produce a substantial variety of 70–130 kDa crystal forming proteins – δ-endotoxins – displaying highly specific entomocidal activity toward larvae of different insect orders. The presence of multiple δ-endotoxin genes (*cry* genes) in the same cell appears to be a characteristic feature of *BT*. Certain strains of subspecies *wuhanensis* and *galleriae* as well as some strains of other subspecies were reported earlier to produce up to 7 different Cry proteins [1,2].

A functional role of the multiple *cryI* genes in the genomes of *BT* and their products in the crystals was discussed. It was proposed to serve as an adaptation to amplify the host range, to prevent resistance development and to increase the activity level due to a cooperative interaction of toxins in the course of targeted membrane pore formation. Mutual expression regulation characteristic of multiple *cry* genes was not investigated sufficiently with respect to either the optimal stoichiometry of the crystalline proteins or to the molecular mechanism providing a different gene activity level.

The complete sequencing of all genes from a certain strain is a prerequisite step to reconstructing the *cry* gene and Cry protein multicomponent complex. This might help to clarify the molecular basis of *cry* gene natural overexpression and the mechanism of toxic action, as well as in designing an improved formula of the available δ-endotoxin combinations for practical use.

2. Materials and methods

2.1. Genomic gene bank construction

Total genomic DNA of *BT wuhanensis* was prepared as described [3], partially digested by *Eco*RI restrictase without purification in agarose gel and ligated with *Eco*RI-linearized λpSL5 phagemid vector. The ligated DNA was packed to λ-phage capsid as described in [3] and applied to an *E. coli* Y1090 (*supF*) lawn grown at 37°C. The primary bank in the phage form contained 10000 colonies. This was washed from the plates, giving up to 10^9 phage particles per μl. It was applied for screening in the phage form at a density of 3000 colonies per plate.

2.2. cryIM cloning and sequence analysis

Screening of the bank was performed using antiserum raised against total crystals of *BT galleriae* 11-67 [4] preliminarily exhausted toward total proteins of *E. coli* Y1090. The clones were rescreened using the same serum and a probe derived from the cloned *cryIAa* gene. Selected phagemid clones were used to transduce *E. coli* TG1 cells. Resultant clones (plasmid form) were cultivated at 30°C and used to isolate phagemid DNA according to a modified alkaline method of Birnboim and Doily [5].

2.3. Sequencing

Sequencing was performed using a manual Sanger dideoxynucleotide method with double-stranded templates and DNA sequenase version 2 [6] or automatically with a thermostable DNA polymerase.

2.4. Expression of the cloned gene and the protein assay

Expression of the genes cloned in pUC derivatives was performed in *E. coli* JM101 and TG1 strains [5]. Cultivation medium (1% peptone, 1% NaCl and 0.5% beef extract) was supplemented with 150 μg ampicillin. The cultures were maintained at 30°C. An overnight culture was diluted 100-fold with fresh medium and grown for 6–10 h. The cells were harvested, treated with lysozyme at a concentration of 10 μg/ml and sonicated. The insoluble cell fraction was collected by centrifugation, than washed twice with water and twice with 30 mM NaOH containing 0.1% Triton X-100. The resulting pellet was suspended in 10 mM EDTA solution and stored at 4°C under a xylene layer.

Proteins of the preparation were assayed by SDS gel electrophoresis and Western blotting according to standard protocols [5].

3. Results and discussion

3.1. cryIM cloning and sequencing

The genomic bank of *BT wuhanensis* 11-76 on the basis of

*Corresponding author. Fax: (7) (095) 315 05 01.

PII S0014-5793(97)00114-2

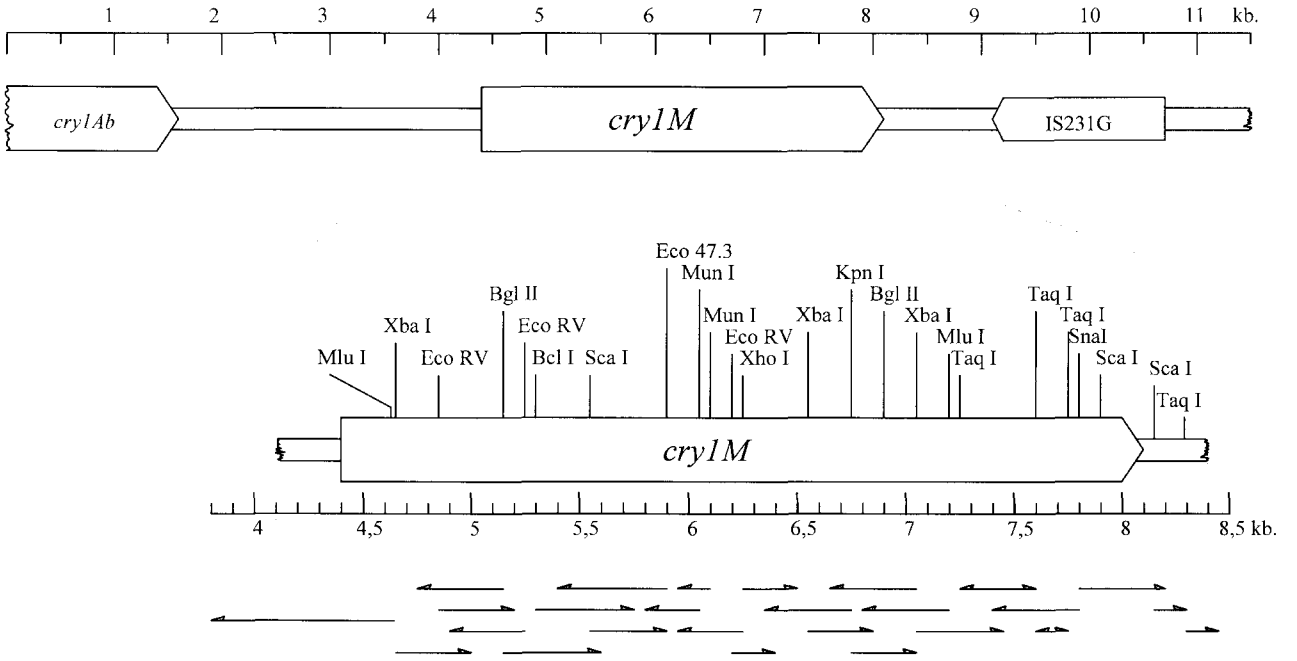

Fig. 1. Structural organization of the *cryIM* containing segment of *Bacillus thuringiensis* ssp. *wuhanensis*. A map and gene location in the 11 kb insertion segment of the λpV7/pVB701 recombinant constructs. The completely sequenced region is shown in a zoom. The locations of the IS231-like sequence and *cryIAb* fragment are predicted by application of a partial sequencing data to the restrictase map.

the λpSL5 phagemid vector contained about 10 000 independent clones with an average insertion length of 12.5 kb (20 genome equivalents). 3000 clones were screened by serum against *BT galleriae* multicomponent crystals and 11 independent clones were selected. They were rescreened with the same antiserum and *cryIAa*-derived probe (520 bp *Hin*dIII fragment from the part encoding a highly conserved and non-toxic C-terminal part of the protoxin). Two clones turned definitely positive in these tests. A recombinant phagemid λpV7 (with 11 kb insertion) was submitted to extensive restrictase mapping (*Hin*dIII and *Xba*I were initially used, then *Bgl*II, *Eco*RV, *Eco*47.III, *Mun*I, *Mlu*I, *Sca*I, *Xho*I, *Kpn*I and *Bst*1107.I were added). Deletion of the λ-phage part by the *Bam*HI site was performed, yielding a pVB701 construct harboring a full-length phagemid insertion fragment. Four additional plasmid subclones in pUKtg131 vector were generated using *Xba*I fragments of the insertion. Partial sequencing of the constructs pVX701, pVX74 and pVX76 containing *Xba*I fragments allowed us to allocate the coding region of a new gene within the cloned fragment (Fig. 1). Moreover, a 2 kb fragment of the *cryIAb* gene (3′ part in the same orientation with *cryIM*) was found in an extremity of the construct. A product of a *cryIAb*-like gene was discovered earlier in crystals of the host *BT* strain [1]. 28 constructs on the basis of pUKtg131 were made. The nucleotide sequences of these subclones overlapping the 4.7 kb segment of interest (Fig. 1) were determined. The resulting sequence has been submitted to the EMBL Data library (accession number Y09326). A *cryIAb* fragment was also partially sequenced (data not shown). An IS231-like sequence was located [7] in the 3′-flanking region of *cryIM*, its terminal repeat being shown on Fig. 2. The complete sequence of this element will be reported elsewhere.

3.2. Analysis of the new gene sequence

Analysis of the new nucleotide sequence revealed an open reading frame with a clear homology with other *cryI*-class genes. The highest percentage of identical residues (about 67%) was found on comparison with CryIAa protein from *BT kurstaki* [8] although it was allocated predominantly to conserved regions of the encoded N-terminal part ('true toxin') [10] and to the C-terminal part, non-essential for toxicity. The percentage of coinciding residues between *cryIM* and *cryIAa* in the conserved blocks was as follows: block 1 – 67%, block 2 – 71%, block 3 – 67%, block 4 – 100%, block 5 – 100% [10]. Other *cryI*-class genes were slightly less similar with the new gene (48–62% of coinciding residues).

Analysis of the *cryIM* 5′-flanking region allowed us to find a BTII-promoter-like sequence [9] (Fig. 2). A BTI-type promoter was not identified in *cryIM* whereas this is considered as being responsible for the natural super-production of CryIAa and other CryI. The lack of a functional BTI may cause a putative natural crypticity of *cryIM*. At the same time, BTII may provide a certain level of expression in some cases [8,9], and thus special experiments were carried out to clarify its functionality.

3.3. Expression of cryIM in E. coli

A pUC19-derived pVB701 construct harboring the full-length 11 kb insertion was applied for expression of *cryIM* in *E. coli*. It contained *cryIM* in the opposite orientation to the *lacZ* promoter and we tried to use the gene's own promoter. Considerable production of a protein with a molecular mass 130 kDa was detected (Fig. 3). It gave a positive reaction with the antiserum used for the bank screening (data not shown). The protein was soluble at pH 9.5, which indicated that the δ-endotoxin was properly folded [8]. This result

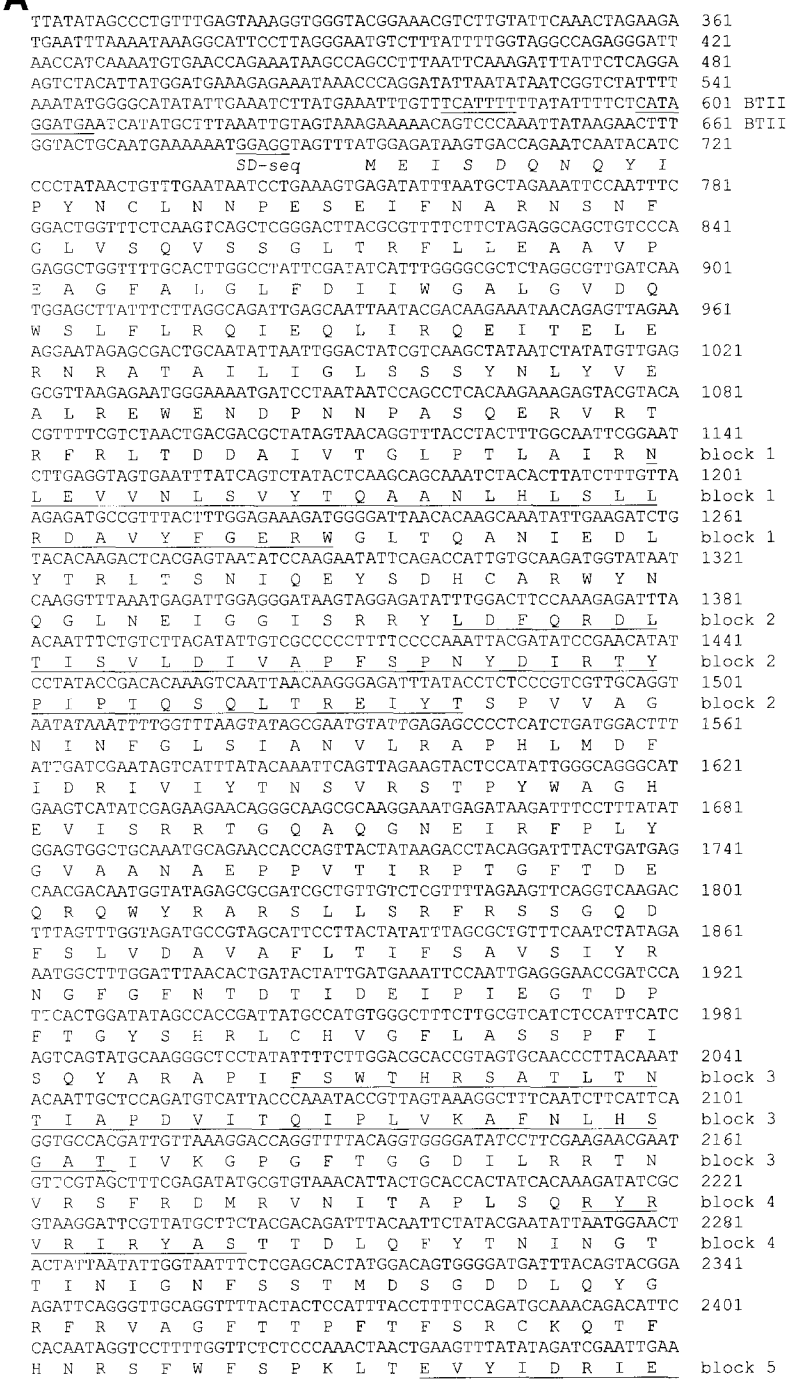

Fig. 2. Nucleotide sequence of *cryIM* gene with translation. Promoter region, translation initiation point and translation, conserved blocks [9,10] and C-terminal processing site are shown by underlining according the alignment. A putative transcriptional terminator is delineated as IR1 and IR2. Short direct and inverted repeats belonging to a putative IS231f-like element are double and single underlined.

proves the functionality of the gene's own *cryIM* promoter in an artificial system described earlier for *cryIAa* [8]. The new gene is the first reported representative of *cry* genes probably possessing only the BTII but not the BTI-type promoter. So far, BTII functionality has not been extensively investigated in either *BT* or *E. coli*. Additional experiments are required to detect its activity in different strains of *BT* and other bacilli.

Acknowledgements: This work was supported by a grant of the Russian Fundamental Research Foundation and a grant of the Russian State program 'New approaches in bioengineering'.

References

[1] Chestukhina, G.G., Kostina, L.I., Zalunin, I.A., Revina, L.P., Mikhailova, A.L. and Stepanov, V.M. (1994) Can. J. Microbiol. 240, 1026–1034.

[2] Baum, J.A. and Gawron-Burke, C. (1991) in: Genetic Engineering, vol. 13 (Setlow, J.K. ed.) pp. 237–263, Plenum, New York.

B

```
TTTGTCCCGGCAGAAGTAACATTTGAGGCAGAATATGATTTAGAGAAAGCTCAGAAAGCG  2521
F  V  P  A  E  V  T  F  E  A  E  Y  D  L  E  K  A  Q  K  A
GTGAATGCGCTGTTTACTTCTTCCAATCAAATCGGGTTAAAAACAGATGTGACGGACTAT  2581
V  N  A  L  F  T  S  S  N  Q  I  G  L  K  T  D  V  T  D  Y
CATATTGATAAAGTATCCAATCTAGTTGAGTGTTTATCAGATGAATTTTGTCTAGATGAA  2641
H  I  D  K  V  S  N  L  V  E  C  L  S  D  E  F  C  L  D  E
AAGCGAGAATTGTCCGAGAAAGTCAAACATGCGAAGCGACTCAGTGATGAGCGGAATTTA  2701
K  R  E  L  S  E  K  V  K  H  A  K  R  L  S  D  E  R  N  L
CTTCAAGATCCAAACTTCAGAGGCATCAATAGACAACCAGACCGTGGTTGGAGAGGAAGT  2761
L  Q  D  P  N  F  R  G  I  N  R  Q  P  D  R  G  W  R  G  S
ACGGATATTACCATCCAAGGAGGAGATGACGTATTCAAAGAGAATTACGTTACGCTACCG  2821
T  D  I  T  I  Q  G  G  D  D  V  F  K  E  N  Y  V  T  L  P
GGTACCTTTGATGGGTGCTATCCAACGTATTTATATCAAAAAATAGATGAGTCGAAATTA  2881
G  T  F  D  G  C  Y  P  T  Y  L  Y  Q  K  I  D  E  S  K  L
AAAGCCTATACCCGTTACCAATTAAGAGGGTATATCGAAGATAGTCAAGACTTAGAAATC  2941
K  A  Y  T  R  Y  Q  L  R  G  Y  I  E  D  S  Q  D  L  E  I
TATTTAATTCGCTACAATGCAAAACACGAAACAGTAAATGTGCCAGGTACGGGTTCCTTA  3001
Y  L  I  R  Y  N  A  K  H  E  T  V  N  V  P  G  T  G  S  L
TGGCCGCTTTCAGCCCAAAGTCCAATCGGAAAGTGTGGAGAGCCGAATCGATGCGCGCCA  3061
W  P  L  S  A  Q  S  P  I  G  K  C  G  E  P  N  R  C  A  P
CACCTTGAATGGAATCCTTACTTAGATTGTTCGTGTAGGAATGGAAAAAAGTGTGCCCAT  3121
H  L  E  W  N  P  Y  L  D  C  S  C  R  N  G  K  K  C  A  H
CATTCGCATCATTTCTCCTTAGACATTGATGTTGGATGTACAGATCTAAATGAGGACCTA  3181
H  S  H  H  F  S  L  D  I  D  V  G  C  T  D  L  N  E  D  L
GGTGTATGGGTGATCTTTAAGATTAAGACGCAAGATGGTCATGCAAGATTAGGAAATCTA  3241
G  V  W  V  I  F  K  I  K  T  Q  D  G  H  A  R  L  G  N  L
GAGTTTCTCGAAGAGAAACCATTATTAGGGGAAGCACTAGCTCGTGTGAAAAGAGCGGAG  3301
E  F  L  E  E  K  P  L  L  G  E  A  L  A  R  V  K  R  A  E
AAAAAATGGAGAGACAAACGTGAAAAATTGGAATTGGAAACAAATATTGTTTATAAAGAG  3361
K  K  W  R  D  K  R  E  K  L  E  L  E  T  N  I  V  Y  K  E
GCAAAAGAATCTGTAGATGCTTTATTTGTAAACTCTCAATATGATCAATTACAAGCGGAT  3421
A  K  E  S  V  D  A  L  F  V  N  S  Q  Y  D  Q  L  Q  A  D
ACGAATATCGCGATGATTCATGCGGCAGATAAACGCGTTCATAGCATTCGAGAAGCGTAT  3481
T  N  I  A  M  I  H  A  A  D  K  R  V  H  S  I  R  E  A  Y
CTGCCGGAGCTGTCTGTGATTCCGGGTGTCAATGCGGCGATTTTTGAAGAATTAGAAGGG  3541
L  P  E  L  S  V  I  P  G  V  N  A  A  I  F  E  E  L  E  G
CGTATTTTCACTGCATTCTCCCTATATGATGCGAGAAATGTCATTAAAAATGGCGATTTC  3601
R  I  F  T  A  F  S  L  Y  D  A  R  N  V  I  K  N  G  D  F
AATAATGGCTTATCATGCTGGAACGTGAAAGGGCATGTAGATGTAGAAGAACAGAACAAC  3661
N  N  G  L  S  C  W  N  V  K  G  H  V  D  V  E  E  Q  N  N
CATCGTTCGGTCCTTGTTGTTCCAGAATGGGAAGCAGAAGTGTCACAAGAAGTTCGTGTT  3721
H  R  S  V  L  V  V  P  E  W  E  A  E  V  S  Q  E  V  R  V
TGTCCGGGTCGTGGTTATATCCTTCGTGTCACAGCGTACAAGGAGGGATATGGAGAAGGT  3781
C  P  G  R  G  Y  I  L  R  V  T  A  Y  K  E  G  Y  G  E  G
TGCGTAACCATTCATGAGTCGAGAACAATACAGCGAACTGAAATTCAGCAACTGCTAGAA  3841
C  V  T  I  H  E  S  R  T  I  Q  R  T  E  I  Q  Q  L  L  E
GAGGAAGTATATCCAAACAACACGGTAACGTGTAATGATTATACTGCAAATCAAGAAGAA  3901
E  E  V  Y  P  N  N  T  V  T  C  N  D  Y  T  A  N  Q  E  E
TACAAAGGTGCGTACACTTCTCATAATCGAGGATATGACGAAGCCTATGGAAATAACCCT  3961
Y  K  G  A  Y  T  S  H  N  R  G  Y  D  E  A  Y  G  N  N  P
TCCGTACCAGCAGATTATACGCCAGTCTATGAAGAAAAAGCGTATACAGATGGACGAAGA  4021
S  V  P  A  D  Y  T  P  V  Y  E  E  K  A  Y  T  D  G  R  R
GACAATCCTTGTGAATCTAACAGAGGGTATGGGGATTACACGCCACTACCAGCTGGTTAT  4081
D  N  P  C  E  S  N  R  G  Y  G  D  Y  T  P  L  P  A  G  Y
GTGAGTAAGGAATTAGAGTACTTCCCAGAAACCGATAAGGTATGGATTGAGATTGGAGAA  4141
V  S  K  E  L  E  Y  F  P  E  T  D  K  V  W  I  E  I  G  E
ACGGAAGGAACATTTATCGTGGAGAGCGTGAATTACTCCTTATGGAGGAATAGGATACGT  4201
T  E  G  T  F  I  V  E  S  V  N  Y  S  L  W  R  N  R  I  R
TTTAAAATGTAACGTGTGCAAATAAAGAATAAGTACTGATCTTTATTAACAGATAAATAA  4261
F  K  M  U
GAAAGTTTATCTGTTAATAAAAAACGCGCATCACTCTTAAGTGAATGATGTCCGTTTTTT  4321
TTATGATTTGATTCAACAAGTGATATGTAAATATATTTTTTTGCGAAAGTTTTACATAAC  4381
AAAAAAATTCGTATAGCAAAATTCTAAATTCAACCTTAAATATCGTTGGGGTGAAAATAT  4441
GCCAAACTAATTTATTCCGAATGTTTAATCGAAACAAATCATAAACAGAAAATACAGGTA  4501
TAAGGGGCACCATACATGCCCATCAACTTAAGGATGGATACAAACAAAATTTTGTTCATT  4561 DR and IR IS231F-type
TGAACTGAAA
```

Fig. 2 (*continued*).

[3] Yankovsky, N.K., Fonstein, M.Yu. and Lashina, S.Y. (1989) Gene 81, 203–207.
[4] Chestukhina, G.G., Kostina, L.I., Zalunin, I.A., Khodova, O.M. and Stepanov, V.M. (1988) FEBS Lett. 232, 249–251.
[5] Sambrook, J., Fritsch, E., Maniatis, T. (1989) Molecular Cloning; 2nd edn., vol. 3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[6] Tabor, S. and Richardson, C.C. (1987) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[7] Rezsohazy, R., Hallet, B. and Delcour, J. (1992) Mol. Microbiol. 6, 1959–1967.
[8] Schnepf, H.E., Wong, H.C. and Whitely, H.R. (1987) J. Bacteriol. 169, 4110–4118.
[9] Brown, K.L. and Whitely, H.R. (1988) Proc. Natl. Acad. Sci. 85, 4166–4170.
[10] Hofte, H. and Whitely, H.R. (1989) Microbiol. Rev. 53, 242–255.

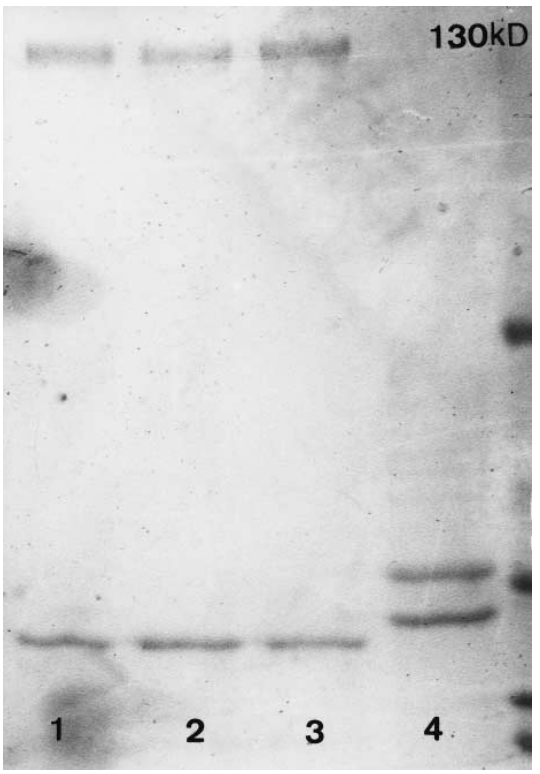

Fig. 3. Expression of *cryIM* in *E. coli*. The pVB701 (*cryIM*) construct and control plasmids pKP7 and pUC18tox harboring *cryIG* and *cryIAa* genes were transformed to TG1 and JM101 strains of *E. coli* and resulted clones were cultivated on a medium without yeast extract (see Section 2) to prevent catabolic repression activation and *lacZ* promoter leakage. The cells were harvested, sonicated and multiply washed with a neutral 10 mM EDTA solution to remove soluble proteins. The insoluble fraction was applied to a denaturative PAAG electrophoresis which was further stained with CBB-R250. (1) pVB701 (*cryIM*); (2) pKP7 (*cryIG*); (3) pUC18tox (*cryIAa*); (4) non-recombinant TG1 strain of *E. coli* (negative control); (5) molecular mass standards (65, 45, 36, 29 and 24 kDa).